Formal Semantics and Pragmatics for Natural Language Querying

James Clifford
(New York University)

Cambridge, England: Cambridge University Press, 1990, xv + 194 pp. (Cambridge tracts in theoretical computer science 8) Hardbound, ISBN 0-521-35433, $29.95

Reviewed by
Alice ter Meulen
Indiana University

Clifford establishes with his semantics of temporal reference a major connection between two areas of computational semantics, namely the semantics of databases and the semantics of natural language. He argues that a semantic representation of time is essential in making database queries more natural and efficient. Chapter 2 presents essentially a variant of Gallin's (1975) TY_2: an intensional logic, IL_s, which is Montague's formal language of PTQ enriched with constants and variables for temporal indices and interpreted with respect to simple reference points consisting only of times—that is, dropping possible worlds and all modality. There are two important differences between IL_s and Montague's Intensional Logic:

- The interpretation function maps a nonlogical constant of type *a* to an element in its extension, instead of to a function from possible worlds to extensions of the appropriate type as in PTQ.
- PTQ's hidden function application of evaluating the intension of a constant, an individual concept, at an index is made explicit by lambda-abstraction over the index and applying that term to the index of evaluation.

The first change avoids the PTQ-style detour via intensional types for interpretations that do not need it; the second crucially affects the account of queries Clifford proposes. The exposition is technical even for a reader initiated in PTQ, but otherwise quite clear, although discussion and motivation follows the presentation of the formal language and its model theory.

In Chapters 3 and 4, the tools for modeling changing states of information about the world are presented: an historical database. Conventional databases represent static snapshots of the world at a given moment in time, and substitution of data loses the old information. An historical database models information about change in the world, and old information can be retained, governed by certain constraints. Facts are recorded with respect to a state, which 'time-stamps' it. Intervals come into play in the Continuity Assumption, which requires that a fact be interpreted as true—continuously from the time it is recorded until a different fact of the same sort is recorded. The inferential differences between the various aspectual classes, for which Dowty (1979) modified PTQ to evaluate formulas with respect to intervals, are said to play no role in the intended historical database application, so expressions are evaluated with respect to a moment in time. Hence the progressive is equivalent to

the simple past tense—all information is stative. Database attributes are functions from moments to appropriate values in extensional domains, and IL_s expresses properties of these higher-order objects by quantifying over moments in time. Intensional queries that use the historical database are, for instance, *Has Peter's salary risen?*, *When was Peter rehired?*, or *Has Liz ever earned the same as Peter?* The meaning of intensional predicates is defined indirectly through meaning postulates; e.g. *rise* is defined as true of something if its current value is higher than the value first recorded for it during its lifespan. Some lexical relations between nouns and verbs are encoded, but lexical decomposition has been avoided. The database creates an object with a limited lifespan representing the salary, as dollar amounts do not rise but salaries do (a variant of the classical temperature puzzle in Montague's PTQ). It should be emphasized that this notion of an historical database is an abstract theoretical concept. Implementation of such rich databases would be prohibitively expensive, as much of the inherent data redundancy cannot be avoided. But a small fragment has been implemented in Prolog by Clifford, as a pilot study, along the lines of Friedman and Warren (1978).

Chapters 5, 6, and 7 present the query language QE-III, the theory of questions underlying it, and a formal fragment with some well-chosen illustrations of the advantages of this approach. Various theories of interrogatives that have attempted to encode some representation of the answer to a question as part of its denotation, e.g. *wh*-questions denoting functions from sets of properties to sets of individuals that have those properties, have led to a considerably complicated semantics (e.g. Karttunen 1977; Hamblin 1973; Belnap 1982). On such a 'single-semantics' approach, a question such as *Who manages John?* denotes the set that contains for each x who manages John, the proposition that x manages John. Instead, Clifford lets questions denote propositional functions just as declarative sentences with free pronouns. The pragmatic component accounts then for the notion of the answer to a question, using both its syntactic and semantic analysis. For instance, *Who manages whom?* is given the same semantic interpretation as *He manages him*, but the pragmatic interpretation of the question is the set of n-tuples that answer it, whereas the declarative is not affected by the pragmatic rules.

Other work on interrogatives shows a similar tendency to defer part of the interpretation to the pragmatics. Clifford briefly mentions Gunji (1981) with his super-interpreter for conversational implicatures, the PHLIQA project reported in Scha (1983), and Hausser and Zaefferer (1978), who encode context-dependency into the semantics by means of context-variables to account for possible redundancy of answers. But no reference at all is made to the important work by Groenendijk and Stokhof (G&S) at the University of Amsterdam, who develop a Montegovian theory of interrogatives in which the semantics accounts for the interpretation of a question also based on Gallin's TY_2, but type-shifting rules in a flexible type theory avoid the semantic complexities that Clifford feared. G&S employ a pragmatic component to account for the relation between the answer and the information already available in the context of use, accounting for the redundancy of a complete, exhaustive list answer in contexts where a property answer or a partial answer may suffice. A deeper and more substantive comparison between Clifford's account and G&S should provide some interesting differences in their empirical and theoretical consequences, and may shed further light on the nature of the pragmatic component and its explanatory power, but such is unfortunately beyond the scope of this book review. Other related research that also deserves mention in this context is Engdahl's (1986) theory of interrogatives in situation semantics, work on hypothetical reasoning in Prolog by, among others, Bonner (1988) and Gabbay (1985), and the event-based PTQ-inspired temporal semantics for Prolog by Richards and Bethke (1989).

Clifford's excellent book is a beautiful example of a foundationally sound and intrinsically interesting computational application of Montague semantics, presented in all requisite detail and with much enlightened discussion of the results obtained, their advantages, disadvantages, and essential limitations. It demonstrates again that the value of theoretical logical research on the semantics of natural language will prove itself in more sophisticated computational applications, offering drastic and fundamental improvements of the existing toolkit.

References

Belnap, Nuel D. Jr. (1982). "Questions and answers in Montague grammar." In S. Peters and E. Saarinen (eds.) *Processes, beliefs, and questions*. Dordrecht: Reidel, 165–198.

Bonner, A.J. (1988). "Hypothetical Datalog: Complexity and expressibility." *Proceedings of the Second International Conference on Database Theory* (ICDT). Springer-Verlag, (Lecture notes in computer science, vol. 326), 144–160.

Dowty, David R. (1979). *Word meaning and Montague grammar*. Dordrecht: Reidel.

Engdahl, Elisabet (1986). *Constituent Questions: The syntax and semantics of questions with special reference to Swedish*. Dordrecht: Reidel.

Friedman, Joyce and Warren, David Scott (1978). "A parsing method for Montague grammars." *Linguistics and Philosophy*, **2**, 347–372.

Gabbay, Dov (1985). "N-Prolog: An extension of Prolog with hypothetical implications." *Journal of Logic Programming*, **2**, 251–283.

Gallin, D. (1975). *Intensional and higher-order modal logic*. Amsterdam: North-Holland.

Groenendijk, J. and Stokhof, M. (1982). "The semantic analysis of *wh*-complements." *Linguistics and Philosophy*, **5**, 175–233.

Groenendijk, J. and Stokhof, M. (1984a). "On the semantics of questions and the pragmatics of answers." In F. Landman and F. Veltman (eds.) *Varieties of formal semantics*. Dordrecht: Foris Publications (GRASS), 143–170.

Groenendijk, J. and Stokhof, M. (1984b) *Studies on the semantics of questions and the pragmatics of answers*. Ph.D. thesis, University of Amsterdam.

Gunji, T. (1981). *Toward a computational theory of pragmatics: Discourse, presupposition, and implicature*. Ph.D. thesis, The Ohio State University.

Hamblin, C.L. (1973). "Questions in Montague English." *Foundations of Language*, **10**(1), 41–53.

Hausser, R. and Zaefferer, D. (1978). "Questions and answers in a context-dependent Montague grammar." In F. Günthner and S. Schmidt (eds.) *Formal semantics and pragmatics for natural languages*. Dordrecht: Reidel, 339–358.

Karttunen, Lauri (1977). "Syntax and semantics of questions." *Linguistics and Philosophy*, **1**, 3–44.

Richards, Barry and Bethke, Unge (1989). "Temporal databases: An IQ approach." In Barry Richards et al. (eds.) *Temporal representation and inference*. London: Academic Press (Cognitive science series). 231–284.

Scha, Remko (1983). *Logical foundations for question answering*. Ph.D. thesis, University of Groningen.

Alice ter Meulen is associate professor of linguistics and philosophy at Indiana University. She has published on mass terms, events and aspect in Montague semantics, and dynamic theories of interpretation. She is co-editor of the GRASS series in natural language semantics (Foris), and recently coauthored with Barbara Hall Partee and Robert Wall the textbook *Mathematical Methods in Linguistics* (Kluwer 1990). Ter Meulen's address is: Department of Philosophy, Sycamore Hall 026, Indiana University, Bloomington, IN 47405; e-mail: atm@iubacs.bitnet